数理解析研究所講究録
第 1821 巻 2013 年 231-238

Optimization for a mixed integer programming problem

新潟大学大学院自然科学研究科　山田修司 (Syuuji YAMADA)
Graduate School of Science and Technology, Niigata University
新潟大学大学院自然科学研究科　田中環 (Tamaki TANAKA)
Graduate School of Science and Technology, Niigata University
大阪大学大学院工学研究科　谷野哲三 (Tetsuzo TANINO)
Graduate School of Engineering, Osaka University

Abstract. In this paper, we propose an interactive solution method for a convex mixed integer programming problem. The algorithm is based on an outer approximation method, a penalty function method and a submodular minimization method. It is shown that the proposed algorithm has the global convergence.

Keywords: Mixed integer programming problem, L$^\natural$-convexity, submodular minimization, outer approximation, penalty function.

1 Introduction

In this paper, we consider a convex mixed integer programming problem (P) with the objective function having the L$^\natural$-convexity. For (P), an outer approximation algorithm has been proposed (see [2]). It is known that an integer programming problem to minimize a L$^\natural$-convex function can be transformed into a submodular minimization problem. For such a problem, a strongly polynomial algorithm has been proposed in [4]. Hence, we propose another outer approximation method for solving (P) by incorporating the submodular minimization algorithm.

The organization of this paper is as follows: In Section 2, we explain L$^\natural$-convexity and a submodular function. In Section 3, we introduce a convex mixed integer programming problem. In Section 4, we describe the outer approximation algorithm proposed by Bonami, Biegler, Conn, Cornuèjols, Grossmann, Laird, Lee, Lodi, Margot, Sawaya and Wächter [2]. In Section 5, we propose another outer approximation algorithm by incorporating a penalty function algorithm and the submodular minimization algorithm proposed by Iwata [4].

2 Mathematical preliminaries

Throughout this paper, we use the following notation: $\mathbb{R}$ and $\mathbb{Z}$ denote the sets of all real numbers and all integer numbers, respectively. Let $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$. For a natural number n, $\mathbb{R}^n$ denotes an n-dimensional Euclidean space. Let $\|\cdot\|$ denote the Euclidean norm. Given a vector $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}^\top$ denotes the transposed vector of $\boldsymbol{x}$. For a vector $\boldsymbol{x} = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$, $\lceil \boldsymbol{x} \rceil$ and $\lfloor \boldsymbol{x} \rfloor$ are the vectors in $\mathbb{Z}^m$ such that the ith elements $\lceil \boldsymbol{x} \rceil_i$ and $\lfloor \boldsymbol{x} \rfloor_i$ are defined as $\lceil \boldsymbol{x} \rceil_i := \min\{\boldsymbol{z} \in \mathbb{Z} : \boldsymbol{z} \geq x_i\}$ and $\lfloor \boldsymbol{x} \rfloor_i := \max\{\boldsymbol{z} \in \mathbb{Z} : \boldsymbol{z} \leq x_i\}$, respectively. Given vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{x} \vee \boldsymbol{y} := (\max\{x_1, y_1\}, \ldots, \max\{x_n, y_n\})^\top$ and $\boldsymbol{x} \wedge \boldsymbol{y} := (\min\{x_1, y_1\}, \ldots, \min\{x_n, y_n\})^\top$. For a function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(\boldsymbol{x})$ denotes the gradient vector of f at $\boldsymbol{x} \in \mathbb{R}^n$. Given a vector $\boldsymbol{x} \in \mathbb{R}^n$ and a positive real number $\varepsilon \in \mathbb{R}$, $B(\boldsymbol{x}, \varepsilon) := \{\boldsymbol{y} \in \mathbb{R}^n : \|\boldsymbol{y} - \boldsymbol{x}\| < \varepsilon\}$. For a subset $D \subset \mathbb{R}^n$, cl D denotes the closer of D.

Moreover, we review some concepts for extended real valued functions.

Definition 2.1 Let S be a nonempty convex set on $\mathbb{R}^n$. A function $p : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* on S if $(1-\lambda)p(\boldsymbol{x}) + \lambda p(\boldsymbol{y}) \geq p((1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{y})$ for each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ $(0 \leq \lambda \leq 1)$.

Definition 2.2 A function $q : \mathbb{Z}^m \to \mathbb{R}$ is said to be L$^\natural$-*convex* if $q(\boldsymbol{a}) + q(\boldsymbol{b}) \geq q\left(\left\lceil \frac{\boldsymbol{a}+\boldsymbol{b}}{2} \right\rceil\right) + q\left(\left\lfloor \frac{\boldsymbol{a}+\boldsymbol{b}}{2} \right\rfloor\right)$ for each $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}^m$.

Lemma 2.1 *Let $q_1, q_2 : \mathbb{Z}^m \to \mathbb{R}$ be L$^\natural$-convex. Then, the following assertions hold.*

(i) $\lambda_1 q_1(\boldsymbol{a}) + \lambda_2 q_2(\boldsymbol{a})$ *is L$^\natural$-convex on $\mathbb{Z}^m$ for each $\lambda_1, \lambda_2 \geq 0$ $(\lambda_1, \lambda_2 \in \mathbb{R})$.*

(ii) $\max\{q_1(\boldsymbol{a}), q_2(\boldsymbol{a})\}$ *is L$^\natural$-convex on $\mathbb{Z}^m$.*

Proposition 2.1 *Let $q : \mathbb{Z}^m \to \mathbb{R}$ be L$^\natural$-convex. Then, there exists a convex function $\bar{q} : \mathbb{R}^m \to \mathbb{R}$ satisfying*

$$\bar{q}(\boldsymbol{a}) = q(\boldsymbol{a}).$$

Definition 2.3 A function $p : \mathbb{R}^n \to \mathbb{R}$ is said to be L$^\natural$-*convex* if $p(\boldsymbol{x}) + p(\boldsymbol{y}) \geq p((\boldsymbol{x} - \alpha \boldsymbol{e}) \vee \boldsymbol{y}) + p(\boldsymbol{x} \wedge (\boldsymbol{y} + \lambda \boldsymbol{e}))$ for each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \geq 0$ $(\lambda \in \mathbb{R})$, where $\boldsymbol{e} := (1, \ldots, 1)^\top \in \mathbb{R}^n$.

Lemma 2.2 *Let $p_1, p_2 : \mathbb{R}^n \to \mathbb{R}$ be L$^\natural$-convex. Then, the following assertions hold.*

(i) $\lambda_1 p_1(\boldsymbol{x}) + \lambda_2 p_2(\boldsymbol{x})$ *is L$^\natural$-convex on $\mathbb{R}^n$ for each $\lambda_1, \lambda_2 \geq 0$ $(\lambda_1, \lambda_2 \in \mathbb{R})$.*

(ii) $\max\{p_1(\boldsymbol{x}), p_2(\boldsymbol{x})\}$ *is L$^\natural$-convex on $\mathbb{R}^n$.*

Proposition 2.2 *Let $p : \mathbb{R}^n \to \mathbb{R}$ be L$^\natural$-convex. Then,*

$$p(\boldsymbol{a}) + p(\boldsymbol{b}) \geq p\left(\left\lceil \frac{\boldsymbol{a}+\boldsymbol{b}}{2} \right\rceil\right) + p\left(\left\lfloor \frac{\boldsymbol{a}+\boldsymbol{b}}{2} \right\rfloor\right)$$

for each $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}^n$.

Definition 2.4 Let V : be a finite set of $\mathbb{R}^n$. Then, a set function $F : 2^V \to \mathbb{R}$ is said to be *submodular* if $F(S) + F(T) \geq F(S \cup T) + F(S \cup T)$ for each $S, T \subset V$.

Definition 2.5 Let $V := \{1, \ldots, m\}$ and let $F : 2^V \to \mathbb{R}$ be submodular. Then, $B(F)$ is said to be the *base polyhedron* of F, where

$$B(F) := \{\boldsymbol{v} = (v_1, \ldots, v_m)^\top \in \mathbb{R}^m : \boldsymbol{v}(V) = F(V),\ \boldsymbol{v}(S) \leq F(S) \text{ for each } S \subset V\},$$

and $\boldsymbol{v}(S) := \sum_{i \in S} v_i$. A vector of $B(F)$ is called a *base*. An extreme point of $B(F)$ is called an *extreme base*.

Proposition 2.3 *Let* $V := \{1,\ldots,m\}$ *let* $F : 2^V \to \mathbb{R}$ *be a submodular function satisfying* $F(\emptyset) = 0$. *Then,*

$$\max\{\boldsymbol{v}^-(V) : \boldsymbol{v} = (v_1,\ldots,v_m)^\top \in B(F)\} = \min\{F(S) : S \subset V\},$$

where $\boldsymbol{v}^-(V) := \sum_{i=1}^{m} \min\{v_i, 0\}$.

Remark 2.1 *For each* $\boldsymbol{v} \in B(F)$,

$$\boldsymbol{v}^-(V) \leq \boldsymbol{v}(V) \leq F(\boldsymbol{v}).$$

Proposition 2.4 *Let* $V := \{1,\ldots,m\}$ *let* $q : \mathbb{Z}^m \to \mathbb{R}$ *satisfy* dom $q \subset \{0,1\}^m$ *and* $F : 2^V \to \mathbb{R}$. *Assume that* $F(S) = q(\chi_S)$ *for each* $S \subset V$, *where* $\chi_S = (\chi_1,\ldots,\chi_m)^\top$ *and*

$$\chi_i := \begin{cases} 1, & \text{if } i \in S, \\ 0, & \text{if } i \notin S. \end{cases}$$

Then, q *is* $\mathrm{L}^\natural$*-convex if and only if* F *is submodular.*

Remark 2.2 *For each* $S, T \subset V$,

$$\left\lceil \frac{\chi_S + \chi_T}{2} \right\rceil = \chi_{S\cup T}, \qquad \left\lfloor \frac{\chi_S + \chi_T}{2} \right\rfloor = \chi_{S\cap T}.$$

3 A mixed integer programming problem

In this paper, we propose a new outer approximation method for solving the following mixed integer programming problem:

$$(\mathrm{P}) \begin{cases} \text{minimize} & f(\boldsymbol{x},\boldsymbol{y}), \\ \text{subject to} & g_i(\boldsymbol{x},\boldsymbol{y}) \leq 0 \quad i = 1,\ldots,l, \\ & \boldsymbol{x} \in \{0,1\}^m, \\ & \boldsymbol{y} \in Y \subset \mathbb{R}^n, \end{cases}$$

where Y is a compact convex set in $\mathbb{R}^n$, $f, g_1, \ldots, g_l : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ are continuous functions. For (P), we assume that the feasible set is nonempty. Since the feasible set is a compact set, (P) has a globally optimal solution.

4 An outer approximation algorithm

In this section, we assume the following conditions for (P).

- Y is a polytope.
- $f(\boldsymbol{x}, \cdot)$ and $g_i(\boldsymbol{x}, \cdot)$ $(i = 1,\ldots,l)$ are continuously twice differentiable convex functions on $\mathbb{R}^n$ for each $\boldsymbol{x} \in \{0,1\}^m$.
- $f(\cdot, \boldsymbol{y})$ and $g_i(\cdot, \boldsymbol{y})$ $(i = 1,\ldots,l)$ are continuously twice differentiable convex functions on $\mathbb{R}^m$ for each $\boldsymbol{y} \in \mathbb{R}^n$.

Moreover, let us consider the following convex programming problem:

$$(\tilde{\mathrm{P}})\begin{cases} \text{minimize} & f(\boldsymbol{x},\boldsymbol{y}),\\ \text{subject to} & g_i(\boldsymbol{x},\boldsymbol{y}) \le 0 \quad i=1,\dots,l,\\ & 0 \le \boldsymbol{x} \le \boldsymbol{e},\ \boldsymbol{x}\in\mathbb{R}^m,\\ & \boldsymbol{y}\in Y. \end{cases}$$

Then, we have the following inequality.

$$\min(\mathrm{P}) \ge \min(\tilde{\mathrm{P}}),$$

where $\min(\mathrm{P})$ and $\min(\tilde{\mathrm{P}})$ denote the optimal values of (P) and $(\tilde{\mathrm{P}})$, respectively. Moreover, for given a finite set $D := \{(\boldsymbol{x}^1,\boldsymbol{y}^1),\dots,(\boldsymbol{x}^k,\boldsymbol{y}^k)\} \subset \mathbb{R}^m\times\mathbb{R}^n$, we consider the following mixed integer programming problem:

$$(\mathrm{P}^{\mathrm{OA}}(D))\begin{cases} \text{minimize} & \alpha\\ \text{subject to} & \nabla f(\boldsymbol{x}^j,\boldsymbol{y}^j)^\top \begin{pmatrix} \boldsymbol{x}-\boldsymbol{x}^j\\ \boldsymbol{y}-\boldsymbol{y}^j\end{pmatrix} + f(\boldsymbol{x}^j,\boldsymbol{y}^j) \le \alpha\\ & \nabla g_1(\boldsymbol{x}^j,\boldsymbol{y}^j)^\top \begin{pmatrix} \boldsymbol{x}-\boldsymbol{x}^j\\ \boldsymbol{y}-\boldsymbol{y}^j\end{pmatrix} + g_1(\boldsymbol{x}^j,\boldsymbol{y}^j) \le 0\\ & \vdots \qquad\qquad\qquad\qquad\qquad\qquad j=1,\dots,k,\\ & \nabla g_l(\boldsymbol{x}^j,\boldsymbol{y}^j)^\top \begin{pmatrix} \boldsymbol{x}-\boldsymbol{x}^j\\ \boldsymbol{y}-\boldsymbol{y}^j\end{pmatrix} + g_l(\boldsymbol{x}^j,\boldsymbol{y}^j) \le 0\\ & \boldsymbol{x}\in\{0,1\}^m,\ \boldsymbol{y}\in Y,\ \alpha\in\mathbb{R}. \end{cases}$$

Then, the following theorem holds.

Theorem 4.1 ([2], Theorem 1) *For all $\bar{\boldsymbol{x}} \in \{0,1\}^m$, if the following problem is feasible, then define $\bar{\boldsymbol{y}}$ to be its optimal solution.*

$$(\mathrm{P}(\bar{\boldsymbol{x}}))\begin{cases} \text{minimize} & f(\bar{\boldsymbol{x}},\boldsymbol{y})\\ \text{subject to} & g_i(\bar{\boldsymbol{x}},\boldsymbol{y}) \le 0 \quad i=1,\dots,l, \quad \boldsymbol{y}\in Y. \end{cases}$$

On the other hand, if $(\mathrm{P}(\bar{\boldsymbol{x}}))$ *is infeasible, then* $\bar{\boldsymbol{y}}$ *is defined as an optimal solution to the following problem:*

$$(\mathrm{PF}(\bar{\boldsymbol{x}}))\begin{cases} \text{minimize} & \displaystyle\sum_{i=1}^{l} u_i\\ \text{subject to} & g_i(\bar{\boldsymbol{x}},\boldsymbol{y}) - u_i \le 0,\ u_i \ge 0 \quad i=1,\dots,l, \quad \boldsymbol{y}\in Y. \end{cases}$$

Let $\hat{D}$ be the set of all such pairs $(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$. Assume that the KKT conditions are satisfied at every optimal solution of $(\mathrm{P}(\bar{\boldsymbol{x}}))$ *(or* $(\mathrm{PF}(\bar{\boldsymbol{x}}))$*), then* (P) *and* $(\mathrm{P}^{\mathrm{OA}}(\hat{D}))$ *have the same optimal value.*

Under Theorem 4.1, the following outer approximation method for solving (P) has been proposed by Bonami, Biegler, Conn, Cornuèjols, Grossmann, Laird, Lee, Lodi, Margot, Sawaya and Wächter [2].

Algorithm OA

Step 0: Set a tolerance $\tau > 0$, $\check{\beta}_1 := +\infty$ and $\hat{\beta}_1 := -\infty$. Calculate an optimal solution $(\boldsymbol{x}^1, \boldsymbol{y}^1)$ of the following problem:

$$\begin{cases} \text{minimize} & f(\boldsymbol{x}, \boldsymbol{y}) \\ \text{subject to} & g_i(\boldsymbol{x}, \boldsymbol{y}) \leq 0 \quad i = 1, \ldots, l, \\ & 0 \leq \boldsymbol{x} \leq \boldsymbol{e}, \ \boldsymbol{y} \in Y. \end{cases}$$

Set $D_1 := \{(\boldsymbol{x}^1, \boldsymbol{y}^1)\}$ and $k := 1$, go to Step 0.

Step 1: If $\check{\beta}_k - \hat{\beta}_k < \tau$, then stop: $(\boldsymbol{x}^k, \boldsymbol{y}^k)$ is an approximate solution of (P). Otherwise, go to Step 2.

Step 2: Calculate an optimal solution $(\hat{\alpha}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ of $(\mathrm{P}^{\mathrm{OA}}(D_k))$. Set $\hat{\beta}_{k+1} := \hat{\alpha}$. Go to Step 3.

Step 3: Set $\boldsymbol{x}^{k+1} := \hat{\boldsymbol{x}}$. If $(\mathrm{P}(\boldsymbol{x}^{k+1}))$ is feasible, then $\check{\beta}_{k+1} := \min\{\check{\beta}_k, f(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1})\}$ where $\boldsymbol{y}^{k+1}$ is an optimal solution of $(\mathrm{P}(\boldsymbol{x}^{k+1}))$. Otherwise, set $\check{\beta}_{k+1} := \check{\beta}_k$ and calculate an optimal solution $\boldsymbol{y}^{k+1}$ of $(\mathrm{PF}(\boldsymbol{x}^{k+1}))$. Go to Step 4.

Step 4: Set $D_{k+1} := D_k \cup \{(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1})\}$, $k \leftarrow k + 1$ and return to Step 1.

5 New outer approximation algorithm

In this section, we suppose the following conditions for (P).

- Y is a compact convex set in $\mathbb{R}^n$ satisfying $\mathbf{0} \in \operatorname{int} Y$ and $Y \subset B(\mathbf{0}, r)$ for some $r > 0$.
- $f(\boldsymbol{x}, \cdot\,)$ and $g_i(\boldsymbol{x}, \cdot\,)$ $(i = 1, \ldots, l)$ are continuously differentiable convex functions on $\mathbb{R}^n$ for each $\boldsymbol{x} \in \{0, 1\}^m$.
- $f(\,\cdot\,, \boldsymbol{y})$ and $g_i(\,\cdot\,, \boldsymbol{y})$ $(i = 1, \ldots, l)$ are continuously differentiable $\mathrm{L}^\natural$-convex functions on $\mathbb{R}^m$ for each $\boldsymbol{y} \in \mathbb{R}^n$.
- A feasible solution $(\boldsymbol{x}', \boldsymbol{y}')$ of (P) is given.

Since $(\mathrm{P}^{\mathrm{OA}}(D_k))$ has the constraint condition $\boldsymbol{x} \in \{0, 1\}^m$, it is difficult to find a globally optimal solution of $(\mathrm{P}^{\mathrm{OA}}(D_k))$. In general, $(\mathrm{P}^{\mathrm{OA}}(D_k))$ is solved by utilizing branch and bound procedures. In this paper we propose another algorithm based on outer approximation by incorporating a penalty function method and a submodular minimization method.

Algorithm NOA

Step 0: Set a penalty parameter $M_1 > 0$, $\gamma > 1$, a tolerance $\tau \geq 0$, $(\boldsymbol{x}^1, \boldsymbol{y}^1) := (\boldsymbol{x}', \boldsymbol{y}')$ and $\beta_1 := f(\boldsymbol{x}^1, \boldsymbol{y}^1)$. Construct a polytope $S_1 := \{(\boldsymbol{y}, \xi) \in \mathbb{R}^n \times \mathbb{R} : (-\boldsymbol{r}, 0) \leq (\boldsymbol{y}, \xi) \leq (\boldsymbol{r}, \bar{r})\}$ where $\boldsymbol{r} := (r, \ldots, r)^\top \in \mathbb{R}^n$ and $\bar{r} := r\sqrt{n+1}$, and calculate the set $V(S_1)$ of all vertices of S_1. For convenience, set $V(S_0) := \emptyset$. Set $k = 1$ and go to Step 1.

Step 1:

Step 1-0: Set $\{(\boldsymbol{z}^1, \xi_1), \ldots, (\boldsymbol{z}^{\rho_k}, \xi_{\rho_k})\} := V(S_k) \backslash V(S_{k-1})$, $\tilde{M} := M_k$, $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) := (\boldsymbol{x}^k, \boldsymbol{y}^k)$, $\tilde{\beta} := \beta_k$ and $p := 1$, where ρ_k is the number of all elements of $V(S_k) \backslash V(S_{k-1})$. Go to Step 1-1.

Step 1-1: Calculate the optimal value $\Gamma(\boldsymbol{z}^p)$ and an optimal solution $\boldsymbol{x}(\boldsymbol{z}^p)$ of the following problem, and go Step 1-2.

$$(\mathrm{SP}(\boldsymbol{z}^p,\tilde{M}))\begin{cases}\text{minimize} & \Phi(\boldsymbol{x},\boldsymbol{z}^p,\tilde{M}) = f(\boldsymbol{x},\boldsymbol{z}^p) + \tilde{M}\sum_{i=1}^{l}\max\{0, g_i(\boldsymbol{x},\boldsymbol{z}^p)\} \\ & \qquad - f(\mathbf{0},\boldsymbol{z}^p) - \tilde{M}\sum_{i=1}^{l}\max\{0, g_i(0,\boldsymbol{z}^p)\} \\ \text{subject to} & \boldsymbol{x}\in\{0,1\}^m.\end{cases}$$

Step 1-2:

- If $\Phi(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p,\tilde{M}) + f(\mathbf{0},\boldsymbol{z}^p) < \tilde{\beta}$ and $g_i(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p) > 0$ for some $i\in\{1,\dots,l\}$, then set $\tilde{M}\leftarrow\gamma\tilde{M}$ and return to Step 1-1.
- If $p<\rho_k$ and $\Phi(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p,\tilde{M}) + f(\mathbf{0},\boldsymbol{z}^p) \geq \tilde{\beta}$, then set $p\leftarrow p+1$ and return to Step 1-1.
- If $p=\rho_k$ and $\Phi(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p,\tilde{M}) + f(\mathbf{0},\boldsymbol{z}^p) \geq \tilde{\beta}$, then go to Step 1-3.
- If $p<\rho_k$ and $g_i(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)\leq 0$ for all $i\in\{1,\dots,l\}$, then set $(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}})\leftarrow(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)$, $\tilde{\beta}\leftarrow f(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)$, $p\leftarrow p+1$, and return to Step 1-2.
- If $p=\rho_k$ and $g_i(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)\leq 0$ for all $i\in\{1,\dots,l\}$, then set $(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}})\leftarrow(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)$, and $\tilde{\beta}\leftarrow f(\boldsymbol{x}(\boldsymbol{z}^p),\boldsymbol{z}^p)$, and go to Step 1-3.

Step 1-3: Set $M_{k+1} := \tilde{M}$, $(\boldsymbol{x}^{k+1},\boldsymbol{y}^{k+1}) := (\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}})$ and $\beta_{k+1} := \tilde{\beta}$. Choose $(\boldsymbol{w}^k,\mu_k)\in V(S_k)$ satisfying $(\boldsymbol{w}^k)^\top\boldsymbol{w}^k + (\mu_k)^2 = \max\{\boldsymbol{y}^\top\boldsymbol{y}+\xi^2 : (\boldsymbol{y},\xi)\in V(S_k)\}$. Go to Step 2.

Step 2: If $(\boldsymbol{w}^k)^\top\boldsymbol{w}^k + (\mu_k)^2 \leq \bar{r}^2+\tau$, then stop: $(\boldsymbol{x}^{k+1},\boldsymbol{y}^{k+1})$ is an approximate solution and β_{k+1} is an approximate value of the optimal value of (P). Otherwise, go to Step 3.

Step 3: Set $S_{k+1} := S_k\cap\{(\boldsymbol{y},\xi)\in\mathbb{R}^n\times\mathbb{R} : (\boldsymbol{w}^k)^\top\boldsymbol{y}+\mu_k\xi\leq\bar{r}^2\}$ and the vertex set $V(S_{k+1})$. Set $k\leftarrow k+1$ and return to Step 1.

At Step 1-1, for each $p\in\{1,\dots,\rho_k\}$, we have

$$f(\boldsymbol{x},\boldsymbol{z}^p) - f(\mathbf{0},\boldsymbol{z}^p) \leq \Phi(\boldsymbol{x},\boldsymbol{z}^p,\tilde{M}) \quad \text{for each } \boldsymbol{x}\in\mathbb{R}^m.$$

Moreover, we note that if $(\boldsymbol{x},\boldsymbol{z}^p)$ is a feasible solution of (P), then

$$f(\boldsymbol{x},\boldsymbol{z}^p) - f(\mathbf{0},\boldsymbol{z}^p) = \Phi(\boldsymbol{x},\boldsymbol{z}^p,\tilde{M}).$$

Furthermore, in the case where $(\hat{\boldsymbol{x}},\boldsymbol{z}^p)$ is a feasible solution of (P) for some $\hat{\boldsymbol{x}}\in\{0,1\}^m$, it is known that there exists an exact penalty parameter $\hat{M}$ for $(\mathrm{SP}(\boldsymbol{z}^p,\tilde{M}))$ (see, for instance, Theorem 9.3.1 in [1]). This implies that for each $\boldsymbol{x}\in\{0,1\}^m$ satisfying $g_i(\boldsymbol{x},\boldsymbol{z}^p)>0$ for some $i\in\{1,\dots,l\}$, there exists $\boldsymbol{x}'\in\{0,1\}^m$ such that $g_i(\boldsymbol{x}',\boldsymbol{z}^p)\leq 0$ for all $i\in\{1,\dots,l\}$ and

$$\Phi(\boldsymbol{x},\boldsymbol{z}^p,\hat{M}) > \Phi(\boldsymbol{x}',\boldsymbol{z}^p,\hat{M}),$$

because $f(\,\cdot\,,\boldsymbol{z}^p)$ and $g_i(\,\cdot\,,\boldsymbol{z}^p)$ $(i=1,\dots,l)$ are convex functions and $\{0,1\}^m$ is compact. Then, we have

$$\begin{aligned}&\min(\mathrm{SP}(\boldsymbol{z}^p,\hat{M}))\\ &= \min\{f(\boldsymbol{x},\boldsymbol{z}^p) - f(\mathbf{0},\boldsymbol{z}^p) : g_i(\boldsymbol{x},\boldsymbol{z}^p)\leq 0 \text{ for all } i=1,\dots,l,\ \boldsymbol{x}\in\{0,1\}^m\}\\ &\geq \min(\mathrm{P}) - f(\mathbf{0},\boldsymbol{z}^p),\end{aligned}$$

where $\min(\mathrm{SP}(z^p, \hat{M}))$ and $\min(\mathrm{P})$ denote the optimal values of $(\mathrm{SP}(z^p, \hat{M}))$ and (P), respectively. Hence, it follows that for each k,

$$f(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1}) = \min\left\{f(\boldsymbol{x}^k, \boldsymbol{y}^k), f(\boldsymbol{x}(\boldsymbol{z}^p), \boldsymbol{z}^p) : p = 1, \ldots, \rho_k\right\} \geq f(\boldsymbol{x}^k, \boldsymbol{y}^k) \geq \min(\mathrm{P}).$$

Moreover, since $(\boldsymbol{x}^1, \boldsymbol{y}^1)$ is a feasible solution of (P), by the procedure for generating $(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1})$ at iteration k of Algorithm NOA, the sequence $\{(\boldsymbol{x}^k, \boldsymbol{y}^k)\}$ is contained in the feasible set of (P). We remember that $f(\,\cdot\,, \boldsymbol{y})$ and $g_i(\,\cdot\,, \boldsymbol{y})$ $(i = 1, \ldots, l)$ are L$^\natural$-convex for each $\boldsymbol{y} \in \mathbb{R}^n$. Hence, it follows from Lemma 2.2 that the objective function of $(\mathrm{SP}(\boldsymbol{z}^p, \tilde{M}))$ is a L$^\natural$-convex function. Moreover, by Proposition 2.4, we notice that $(\mathrm{SP}(\boldsymbol{z}^p, \tilde{M}))$ can be reformulated a submodular minimization problem. Therefore, we can obtain $\boldsymbol{x}(\boldsymbol{z}^p)$ by utilizing the strongly polynomial time algorithm proposed by Iwata [4]. Since $(Y \times \mathbb{R}_+) \cap \mathrm{cl}\, B(\mathbf{0}, \bar{r}) \subset S_1$ and $\mathrm{cl}\, B(\mathbf{0}, \bar{r}) \subset \{(\boldsymbol{y}, \xi) \in \mathbb{R}^n \times \mathbb{R} : (\boldsymbol{w}^k)^\top \boldsymbol{y} + \mu_k \xi \leq \bar{r}^2\}$ for each k, we have

$$S_1 \supset S_2 \supset \cdots \supset S_k \supset \cdots \supset (Y \times \mathbb{R}_+) \cap \mathrm{cl}\, B(\mathbf{0}, \bar{r}).$$

We note that $(\boldsymbol{w}^k, \mu_k) \notin S_{k+1}$ if $(\boldsymbol{w}^k, \mu_k)$ does not satisfy the stopping criterion at Step 2. Moreover, the following theorems hold.

Theorem 5.1 *Assume that* $\tau = 0$. *Then, the infinite sequence* $\{(\boldsymbol{w}^k, \mu_k)\}$ *generated by Algorithm* NOA *satisfies* $\lim_{k\to\infty} (\boldsymbol{w}^k)^\top \boldsymbol{w}^k + (\mu_k)^2 = \bar{r}^2$.

Theorem 5.2 *Assume that* $\tau = 0$. *Let* $\boldsymbol{y} \in Y$ *and* $\varepsilon > 0$. *Then, there exists* $k' > 0$ *such that*

$$B\left(\left(\boldsymbol{y}, \sqrt{\bar{r}^2 - \boldsymbol{y}^\top \boldsymbol{y}}\right), \varepsilon\right) \cap V(S_k) \neq \emptyset \text{ for each } k \geq k'.$$

Corollary 5.1 *Assume that* $\tau = 0$. *Let* $\{(\boldsymbol{x}^k, \boldsymbol{y}^k)\}$ *be the infinite sequence of the provisional solutions generated by Algorithm* NOA. *Then, we have*

$$\lim_{k\to\infty} f(\boldsymbol{x}^k, \boldsymbol{y}^k) = \min(\mathrm{P}).$$

From Theorem 5.1, by setting τ as a positive number, Algorithm NOA terminates within a finite number of iterations. Moreover, by Corollary 5.1, we note that Algorithm NOA has the global convergence, that is, every accumulation point of the generated by Algorithm NOA is a globally optimal solution of (P).

6 Conclusions

In this paper, we have proposed an outer approximation algorithm for solving a mixed integer programming problem. The proposed algorithm approximates the set lifted the feasible region of continuous variables on a hemisphere in $\mathbb{R}^{n+1}$ by the sequence of polytopes. By utilizing a penalty function method, the subproblem solved at each iteration can be formulated as a problem to minimize a L$^\natural$-convex function over $\{0,1\}^m$. Hence, an optimal solution of the subproblem can be obtained by using the submodular minimization algorithm proposed by Iwata [4]. It is shown that the proposed algorithm has the global convergence. Moreover, we note that the proposed algorithm is useful in the case where the dimensions of the discrete and continuous variable regions are large and small respectively.

References

[1] Bazaraa, M.S., H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed., John Wiley, New York, 2006.

[2] Bonami, P., Biegler, L.T., Conn, A.R., Cornuéjlos, G., Grossmann, I.E., Laird, C.D., Lee, J., Lodi, J., Margot, F., Sawaya, N., and Wächter, A., "An algorithmic framework for convex mixed integer nonlinear programs," *Discrete Optimization* 5, pp.186–204, 2008.

[3] Edmonds, J., "Submodular functions, matroids, and certain polyhedra," *Combinational Structures and Their Applications*, R. Guy, H. Hanai, N. Sauer and J. Schönheim, eds., Gordon and Breach, pp.69–87, 1970.

[4] Iwata, S., "A fully combinational algorithm for submodular function minimization," *Journal of Combinational Theory, Series B* 84, pp.203–212, 2002.